Hydrol. Earth Syst. Sci., 15, 913–930, 2011
www.hydrol-earth-syst-sci.net/15/913/2011/
doi:10.5194/hess-15-913-2011

Hydrology and Earth System Sciences

Improving catchment discharge predictions by inferring flow route contributions from a nested-scale monitoring and model setup

Y. van der Velde[1,2], J. C. Rozemeijer[3], G. H. de Rooij[4], F. C. van Geer[5,6], P. J. J. F. Torfs[2], and P. G. B. de Louw[3]

[1]Soil Physics, Ecohydrology and Groundwater management Group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands
[2]Hydrology and quantitative water management Group, Wageningen University, P.O. Box 47, 6700 AA Wageningen, The Netherlands
[3]Deltares, P.O. Box 85467, 3508 AL Utrecht, The Netherlands
[4]Department of Soil Physics, Helmholtz Centre for Environmental Research – UFZ, Theodor-Lieser-Strasse 4, 06120 Halle, Germany
[5]Department of Physical Geography, Utrecht University, P.O. Box 80115, 3508 TC Utrecht, The Netherlands
[6]TNO Geological Survey of The Netherlands, P.O. Box 80015, 3508 TA Utrecht, The Netherlands

Received: 8 October 2010 – Published in Hydrol. Earth Syst. Sci. Discuss.: 26 October 2010
Revised: 16 February 2011 – Accepted: 17 February 2011 – Published: 15 March 2011

Abstract. Identifying effective measures to reduce nutrient loads of headwaters in lowland catchments requires a thorough understanding of flow routes of water and nutrients. In this paper we assess the value of nested-scale discharge and groundwater level measurements for the estimation of flow route volumes and for predictions of catchment discharge. In order to relate field-site measurements to the catchment-scale an upscaling approach is introduced that assumes that scale differences in flow route fluxes originate from differences in the relationship between groundwater storage and the spatial structure of the groundwater table. This relationship is characterized by the Groundwater Depth Distribution (GDD) curve that relates spatial variation in groundwater depths to the average groundwater depth. The GDD-curve was measured for a single field site (0.009 km^2) and simple process descriptions were applied to relate groundwater levels to flow route discharges. This parsimonious model could accurately describe observed storage, tube drain discharge, overland flow and groundwater flow simultaneously with Nash-Sutcliff coefficients exceeding 0.8. A probabilistic Monte Carlo approach was applied to upscale field-site measurements to catchment scales by inferring scale-specific GDD-curves from the hydrographs of two nested catchments (0.4 and 6.5 km^2). The estimated contribution of tube drain effluent (a dominant source for nitrates) decreased with increasing scale from 76–79% at the field-site to 34–61% and 25–50% for both catchment scales. These results were validated by demonstrating that a model conditioned on nested-scale measurements improves simulations of nitrate loads and predictions of extreme discharges during validation periods compared to a model that was conditioned on catchment discharge only.

Correspondence to: Y. van der Velde (ype.vandervelde@wur.nl)

1 Introduction

Intensive agriculture in lowland catchments often leads to high nutrient losses and eutrophication of downstream waters (Oenema et al., 2007; Van der Molen, 1998; Vitousek et al., 2009). To identify effective measures to reduce these nutrient loads, the flow routes of water that enter a stream and their nutrient concentrations need to be quantified (Tiemeyer et al., 2010). In densely drained lowland catchments, surface water discharge is fed by groundwater flow toward streams and ditches, tube drain effluent, and overland flow. Many field-scale studies identified tube drain effluent as the major source of nitrate (Tiemeyer et al., 2006; Nangia et al., 2010; Rozemeijer et al., 2010c). However, the field scale at which these contributions can be directly measured (De Vos et al., 2000; Van der Velde et al., 2010a) often is not the scale of interest to water management authorities. Extrapolation of fields site results to entire catchments can easily lead to wrong conclusions as field sites can prove non-representative of the patterns and processes that emerge at larger scales (Sivapalan, 2003; Soulsby et al., 2006; Didszun and Uhlenbrook, 2008). Therefore, our challenge is

to effectively integrate information from field-scale measurements into the prediction of catchment-scale flow route contributions.

In Van der Velde et al. (2010a), we presented the results of a field-scale measurement setup that separated tile drain flow from overland flow and groundwater flow. We also measured discharges at two larger nested scales and showed that, rather than the actual measured volumes at the field site, the characteristic response of individual flow routes can be used to upscale the field-site flow routes to the catchment scale. This elementary upscaling approach was purely based on measured data. A model framework was needed to upscale the measured field-scale fluxes to the catchment-scale for periods without complete sets of measurements.

To develop such a model upscaling approach, Sivapalan (2003) advocated the search for concepts that "easily connect scales, and that can also be easily scaled". This should lead to "a watershed-scale representation that is clearly tied to process descriptions at a lower level of scale, and which is not overly complex". In sloped terrain, scaling research has focused on the way in which hillslopes connect to headwaters (Uchida et al., 2005, Jensco et al., 2009; Tetzlaff et al., 2008; Clark et al. (2009) and headwaters to entire basins (Shaman et al., 2004). Rodgers et al. (2005), Tetzlaff et al. (2007), and Didszun and Uhlenbrook (2008) studied the scaling behavior of both discharge and tracers across nested-scale catchments and found that scaling effects in discharge and solutes could largely be attributed to scale-related morphologic, topographic and land-use features. In contrast to sloped catchments, lowland catchments generally have little morphological heterogeneity and the main flow routes occur at all scales. Therefore, the scale effects in discharge of lowland catchments are primarily driven by scale-differences in drainage density of ditches and tube drains, micro-topography (Appels et al., 2011), and soil type.

Tracer based hydrograph separation is an often applied technique to identify scale effects in contributions of different flow routes to catchment discharge (Didszun and Uhlenbrook, 2008; Tiemeyer et al., 2008; Van der Velde et al., 2010). Tiemeyer et al. (2008) and Rozemeijer et al. (2010c) demonstrated that for tile drained lowland systems, especially nitrate concentrations differ strongly between the major flow routes. Although they also showed that flow route nitrate concentrations may vary over time and uncertainties are substantial, the clear concentration differences make rough estimates of the contribution of individual flow routes to catchment discharge possible. Inversely, catchment nitrate concentration dynamics can be an important verification for simulated flow route contributions in a hydrological model of a lowland catchment ("getting the right answers for the right reasons", Kirchner, 2006).

Upscaling approaches for solute transport often focus on travel time distributions (e.g. Botter et al., 2009 and Van der Velde et al., 2010b). Travel time distributions can be defined at all scales and implicitly account for flow routes by combining the spectra of travel times contributing to each flow route in an overall travel time distribution. However, both Botter et al. (2009) and Van der Velde et al. (2010b) showed that a flow simulation with accurate contributions of flow routes is paramount for representing the dynamics in travel time distributions that is needed to describe solute transport dynamics. In Van der Velde et al. (2010b) we used a spatially distributed groundwater model to upscale field-scale flow route contributions to catchment scales. However, because of very long calculation times and the large number of parameters, groundwater models are inflexible and often inaccurate in simulating fluxes of specific flow routes at catchment scales. Hence, Van der Velde et al. (2009) proposed a new upscaling approach for hydrology in lowland catchments (from here on called the Lowland Groundwater-Surface water Interaction model, LGSI-model). We assumed that each flow route (i.e., ditch and stream drainage, overland flow, and tube drain flow) starts to discharge if the groundwater level exceeds a flow route-specific threshold groundwater level at that location, and that the magnitude of the flux depends on the groundwater level. The contribution of a flow route to the total catchment discharge is calculated by integration over all groundwater levels in the catchment, described by a groundwater depth distribution. Van der Velde et al. (2009) showed that each storage volume of groundwater in the saturated zone corresponds to a unique groundwater depth distribution. They also showed that the relation between storage and groundwater depth distribution can be defined at any spatial scale and thus satisfies Sivapalan's (2003) criterion: it "easily connect scales, and can also be easily scaled".

However, in order to measure this relationship between storage and the groundwater depth distribution at catchment scales relevant to water management authorities, many groundwater depth time series are needed throughout the catchment. This makes this approach laborious and in our previous paper (Van der Velde et al., 2009) we had to resort to spatially distributed transient groundwater modeling to derive this relationship. A workable alternative would be to have a dense network of groundwater monitoring wells on a small area within the catchment, and observe the groundwater levels frequently for a limited time period. Obviously some sort of upscaling is then needed to use this data to characterize the behavior of the entire catchment. We introduce here a nested-scale model setup combined with a probabilistic Monte Carlo approach to achieve this. This approach is tested by evaluating the model uncertainty in discharge predictions during an 8 year period and by evaluating uncertainty in simulated nitrate concentrations dynamics during a month with three successive storm events.

The objectives of this paper are twofold. Firstly, we test whether the LGSI-model can accurately describe all individual flow route fluxes at the field scale. This would increase our confidence in the ability of the LGSI-model to simulate flow route fluxes accurately at the catchment scale where

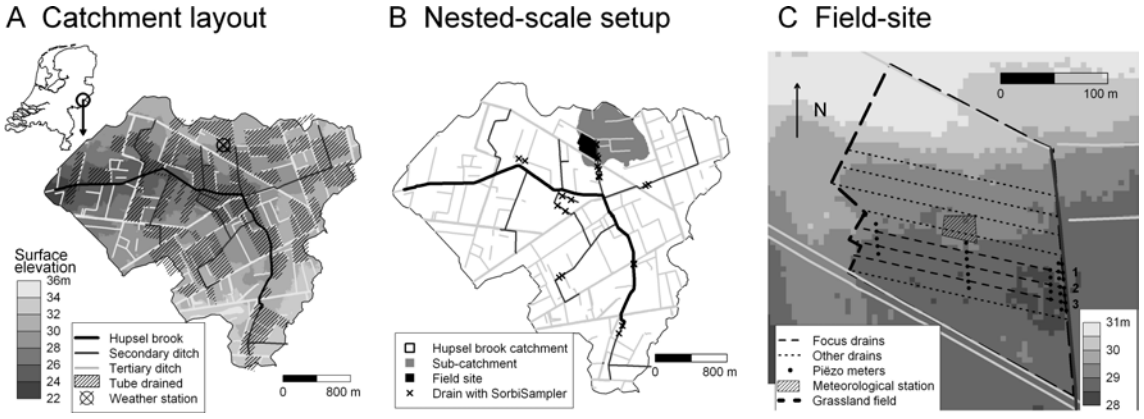

Fig. 1. Hupsel Brook catchment and nested-scale measurement setup.

these fluxes cannot be measured directly. Secondly, we want to assess the value of nested-scale monitoring as presented in Van der Velde et al. (2010a) for reducing uncertainty in predictions of flow route discharges at the catchment scale.

2 Materials and methods

This paper combines the nested-scale measurements introduced by Van der Velde et al. (2010a) and the upscaling approach described in Van der Velde et al. (2009). Therefore, we offer a brief summary of the relevant information (Sects. 2.1 and 2.2), and refer to both papers for detailed background information. The LGSI-model is first applied to the field-site discharge and groundwater level measurements, which is described in Sect. 2.3. Section 2.4 introduces a catchment model conditioned on catchment discharge and a groundwater level time series and a nested-scales model that combines the field-site model and the catchment model. Section 2.5 introduces validation strategies for both the catchment model and the nested-scales model to assess the value of nested scale monitoring.

2.1 Nested experimental setup

The measurements for this study were performed in the Hupsel Brook catchment in the eastern part of The Netherlands (Fig. 1a) (52°06′N; 6°65′E). The size of the catchment is 6.5 km^2, with surface elevations ranging from 22 to 36 m above sea level. At depths ranging from 0.5 to 20 m an impermeable marine clay layer is found (Van Ommen et al., 1989). The unconfined aquifer consists of Pleistocene aeolian sands with occasional layers of clay, peat, and gravel (see Wösten et al. (1985) for more details).

The Hupsel Brook catchment is drained by a straightened and deepened main brook and by a dense artificial drainage network of ditches and tube drains. The spacing between the ditches averages 300 m (Fig. 1a) and approximately 50% of the area is tube drained (plastic perforated flexible tubes). The Hupsel Brook catchment has a semi-humid sea climate with a yearly precipitation of 500 to 1100 mm and a yearly estimated evaporation of 300 to 600 mm.

Within the Hupsel Brook catchment, discharge was measured at three nested spatial scales: (1) the entire catchment of approximately 6.5 km^2, (2) a sub-catchment of 0.4 km^2 and (3) a 0.009 km^2 field site located within the sub-catchment (Fig. 1b). From August 2007 through December 2008, discharge was measured every 15 min for both catchment scales. Continuous surface water nitrate concentrations were measured at the outlet of the entire catchment with a Hydrion-10 multi-parameter probe (Hydrion BV Wageningen, the Netherlands). Monthly average nitrate concentrations of tube drain effluent were measured at 20 locations in the catchment with Sorbi-Samplers (Rozemeijer et al., 2010a).

The tube drained field site of 0.9 ha has a drain spacing of 14.5 m. Along a 43.5 m stretch inside the deep easterly ditch (Fig. 1c), we built in-stream reservoirs with separate vessels to capture tube drain discharge. The in-stream reservoirs collected overland flow and groundwater influx through the stream bed. Thus we separated the tube drain flow from the combined flux of overland flow and groundwater flow. The discharge of both flow routes was measured with 5 min intervals for November 2007 through December 2008. During that period we also manually measured phreatic groundwater levels at 31 locations within the field-site every week. Pressure sensors in 15 piezometers along drain 1 (Fig. 1c) recorded phreatic levels every 10 min. A meteorological station of the KNMI (Royal Dutch Meteorological Institute) bordering the field-site measured hourly rainfall and evapotransipration derived with the Makkink relation (Makkink, 1957).

Table 1. LGSI-Model basics and parameters. The process formulations are point-scale model equations. The catchment-scale equations are obtained by integration over the groundwater depth distribution within appropriate integration bounds (see Appendix B).

Process	Formulation	Process-specific parameters
Unsaturated zone storage (s_{unsat})	$s_{\mathrm{unsat}} = \theta_{\mathrm{s}} \int_0^u \left[1+(\alpha h)^n\right]^{\frac{1}{n}-1} dh$ for $u > 0$ $s_{\mathrm{unsat}} = 0$ for $u < 0$	u: Groundwater depth [L] θ_{s}: Porosity [−] α, n: Van Genuchten parameters [L], [−] h: Height above water table [L]
Saturated storage (s_{surf})	$s_{\mathrm{surf}} = -m \cdot u$ for $u < 0$ $s_{\mathrm{surf}} = 0$ for $u > 0$	m: Fraction of ponding [−]
Evapo-transpiration (e_{act})	$e_{\mathrm{act}} = e_{\mathrm{pot}}$ for $u < u_{\mathrm{ET}}$ $e_{\mathrm{act}} = 0$ for $u > u_{\mathrm{ET}}$	e_{pot}: Potential evapotranpiration [L·T^{-1}] u_{ET}: Evapotranpiration reduction depth [L]
Overland (q_{ov}), Groundwater flow (q_{grw})	$q_{\mathrm{grw,ov}} = \frac{(m-1)\cdot u}{r_{\mathrm{ex}}}$ for $u < 0$ $q_{\mathrm{grw,ov}} = 0$ for $u > 0$	r_{ex}: Exfiltration resistance [T]
Tube drain flow (q_{dr})	$q_{dr} = \frac{D_{dr}-u}{r_{dr}}$ for $u < D_{dr}$ $q_{dr} = 0$ for $u > D_{dr}$	D_{dr}: Tube drain depth [L] r_{dr}: Tube drain resistance [T]
Scale	**Formulation**	**Scale-specific parameters**
GDD-curve	$\sigma_{\mathrm{u}} = \sigma_{\mathrm{diff}} \cdot e^{-\left(\frac{\langle u\rangle - u_{\mathrm{s\ max}}}{b}\right)^2} + \sigma_{\mathrm{min}}$	σ_{u}: Spatial groundwater depth st. dev. [L] $\langle u \rangle$: Spatial average grw. depth [L] σ_{min}: Minimal grw. depth st. dev. [L] σ_{diff}: Maximum increase in grw. depth st. dev. [L] $u_{\mathrm{s\ max}}$: Average grw. depth with maximum grw. depth st. dev. [L]
Surfaces	–	A_{tot}: Catchment area [L^2] A_{dr}: Area with tube drainage [−] A_{s}: Area with surface water [−]

2.2 Lowland Groundwater-surface water interaction (LGSI) model

2.2.1 Model concepts

The LGSI-model (Van der Velde et al., 2009) essentially consists of point-scale expressions of flow route fluxes (tube drain flow, overland flow, groundwater flow, direct rainfall, and evapotranspiration) and storages (saturated storage, unsaturated storage, and surface storage). A point in the catchment starts to generate a flux for a certain flow route when its groundwater level exceeds a threshold specific to that flow route. The magnitude of this flux (except for evapotranspiration) is directly proportional to the difference between the groundwater level and the threshold level. Upscaling of fluxes and storages is achieved by integrating the point-scale expressions over all groundwater depths within a model area. This distribution of groundwater depths was found to approximate a normal distribution. Van der Velde at al. (2009) showed that because of the local nature of groundwater flows towards the nearest ditch or depression in lowland catchments, the overall pattern is a repetition of similarly shaped water tables within fields. The shape within a field depends on the soil type, distance between ditches and drains, surface elevation and the stored water volume. Hence, the catchment-scale distribution of groundwater depths is a summation of the distributions within all fields which approaches a normal distribution via the central limit theorem.

Table 2. Estimated Process-specific and Scale-specific parameter ranges.

Process-specific parameters		Scale-specific parameters			
			Field site	Sub-catchment	Catchment
α	1–2 m[(1)]	$\sigma_{\min}$	0.06–0.13 m[(2)]	0.14–0.22 m[(2)]	0.2–0.30 m[(2)]
n	1–6[1]	σ_{diff}	0.1–0.6 m[(3)]	0.1–0.6 m[(3)]	0.1–0.6 m[(3)]
θ_s	0.35–0.45[(1)]	$u_{s\,\max}$	0.1–0.6 m[(3)]	0.1–0.6 m[(3)]	0.1–0.6 m[(3)]
M	0.05–0.7[(3)]	B	0.1–0.6 m[(3)]	0.1–0.6 m[(3)]	0.1–0.6 m[(3)]
u_{ET}	1–2 m[(3)]	A_{dr}/A_{tot}	1.0	0.7–0.9[(4)]	0.4–0.6[(4)]
r_{ex}	0.1–10 d[(3)]	A_s/A_{tot}	0	0.0054–0.0066[(4)]	0.009–0.011[(4)]
r_{ditch}	500–4000 d[(3)]	A_{tot}	7700–9000 m^2	0.36–0.48 Km^2 [(4)]	6.0–7.3 Km^2 [(4)]
u_{ditch}	1.05 m[(5)]	[(1)] Soil parameter estimates from Wösten et al. (2001)			
D_{dr}	0.75–0.95 m[(5)]	[(2)] DEM			
r_{dr}	100–300 d[(3)]	[(3)] Rough estimates			
a_{dr}	0–2.2[(3)]	[(4)] Topographic maps and field survey			
b_{dr}	0–0.14 d^{-1} [(3)]	[(5)] Field-site measurements			
c_{dr}	0–0.8 mm day^{-1} [(3)]				
L_f	0–0.8 mm day^{-1} [(3)]				
Mp	0.95–1.05[(3)]				
Me	0.95–1.05[(3)]				

This reasoning is valid as long as the scale of the features that create the shape of the water table is far smaller than the model area. The catchment-scale relation between storage and the normal distribution was formalized by the Groundwater Depth Distribution curve (GDD-curve) that describes the relationship between the spatial average groundwater depth and the spatial groundwater depth standard deviation. In summary: saturated storage is calculated by a water balance. Each storage volume corresponds to a distribution of groundwater depths via the GDD-curve. From this distribution all fluxes and unsaturated and surface storage are calculated through simple threshold expressions that relate a water table to a flux or storage.

The model parameters of LGSI-model can be subdivided into process-specific parameters that describe fluxes and storages as a function of the local groundwater level and scale-specific parameters that describe the spatial distribution of groundwater depths, the total catchment area, the tube-drained area, and the area occupied by the surface water network. In Table 1 all point-scale process formulations, the GDD-curve, and their parameters are introduced. Appendix A lists the abbreviations and variables used in this paper. The complete set of LGSI-model equations is summarized in Appendix B. The LGSI-model is a fast calculating process model that calculates flow route discharges for a decade on hourly basis within a few seconds. This is a huge advantage over fully distributed models (e.g. Rozemeijer et al., 2010b), and allows for extensive parameter estimation by Monte Carlo simulation as will be demonstrated in this paper. However, the model is less suited to evaluate the effects of measures that affect the shape of the groundwater table, since this shape is derived from measurements and not calculated from physical principles.

2.2.2 Model extensions for the field site

To apply the LGSI-model to the field site, the basic setup needed to be extended to explicitly include groundwater flow out of the field into the deep ditch, lateral groundwater flow into the field from adjacent fields, and a time-variant flow resistance of the tube drains.

The effects of the single deep ditch to the east of the field on the water table within the field are large compared to the field-site area and hence are not well represented by the assumption of a normal distribution of groundwater depths. To account for groundwater flow towards this ditch, we introduce a new discharge term that approximates the groundwater flux to the deep ditch as a function of the average groundwater depth in the field, $\langle u(t)\rangle$ [L]:

$$Q_{\text{grw,field}}(t) = \frac{u_{\text{ditch}} - \langle u(t)\rangle}{r_{\text{ditch}}} \text{ for } u_{\text{ditch}} > \langle u(t)\rangle \quad (1)$$

with u_{ditch} [L] the depth of the ditch relative to the mean surface elevation of the field site and r_{ditch} [T], the resistance of the field-site to groundwater flow towards the ditch. The lateral groundwater inflow, L_f [LT^{-1}], was assumed constant throughout the simulation period.

The total discharge measured by the in-stream reservoirs of the field experiment, Q_{res}[LT^{-1}], can now be calculated

Table 3. The cutoff criteria for behavioral model runs. Model runs are assigned behavioral when they meet to all "goodness of fit" criteria. Expressions for the error terms are given in Appendix C.

	Field site (storage)			Field site (fluxes)			
	GDD-curve	Ponding curve	Grw. depth	Tube drain	Reservoir	Sub-catchment	Catchment
Curve Error: CE	0.07	0.2	–	–	–	–	–
Cumulative discharge error: CDE	–	–	–	$< 5\%$	$< 10\%$	$< 8\%$	$< 8\%$
Nash-Sutcliff coeff. for time series: *NS*	–	–	> 0.9	> 0.8	> 0.8	> 0.85	> 0.75
Average Groundwater depth error: GE	–	–	–	–	–	< 1 cm (0.2–0.8 p)*	< 1 cm (0.1–0.9 p)*

* The lower and upper quantile of the modeled groundwater depth distribution that is assumed to envelope the measured groundwater depth at the field site. The GE gives the maximum average difference between the measured groundwater depth and the modeled envelope.

by all the water that enters the surface water except for the tube drain flux:

$$Q_{\text{res}}(t) = Q_{\text{grw,field}}(t) + Q_{\text{ov}}(t) + P_{\text{Q}}(t) - E_{\text{Q}}(t) \quad (2)$$

with Q_{ov} the flux by overland flow, P_{Q} rainfall on ponded surfaces (including the ditch) and E_{Q} evaporation from ponded surfaces (all LT^{-1}).

During the experimental period, we observed a strong decline in the drainage effectiveness of the tube drains. At the beginning of the experiment the tube drains were cleaned by pressure flushing as is common practice in the Hupsel Brook catchment. This pressure flushing is repeated every two years. We hypothesized that the tube drains slowly get clogged in periods with substantial discharge and that in dry periods without discharge, aeration and oxidation of the clogging material inside the tube drains reduces the resistance. Similar behavior was also observed by Bentley and Skaggs (1993). The following simple empirical relation was adopted to account for the tube drain resistance change as a function of discharge:

$$\frac{dr_{dr}}{dt} = 1_{\{Q_{dr} > c_{dr}\}} a_{dr} - 1_{\{Q_{dr} < c_{dr}\}} b_{dr} r_{dr} \quad (3)$$

with a_{dr} [–] the rate with which the drainage resistance, r_{dr} [T], increases when the tube drain discharge is larger than threshold discharge c_{dr} [L^3T^{-1}]. The resistance decreases with fractional rate b_{dr} [T^{-1}] for discharges smaller than c_{dr}.

2.2.3 Probabilistic parameter estimation

A parsimonious process model as the LGSI-model necessarily suffers from equifinality (parameter non-uniqueness; Beven and Freer, 2001) stemming from parameter uncertainty, the lumped nature of the parameters, the subjectivity introduced by including and excluding processes, the chosen process formulations, and the many different types of measurements that the model needs to describe. We dealt with equifinality by generating many combinations of parameters in a Monte Carlo procedure (GLUE, Beven and Freer, 2001).

In this study we introduce three LGSI-models: a field-site model, a catchment model, and a nested-scales model. For each model, random parameter values were generated from prior uniform and independent distributions between predetermined parameter ranges. All parameter ranges are listed in Table 2. Parameter sets were qualified behavioral when the model satisfactory described the measured data and all behavioral parameter sets were considered equally probable. The criteria that divide the parameter space in behavioral and non-behavioral parameter sets for each of the models are listed in Table 3. This procedure was continued for each model until an ensemble of 500 behavioral parameter sets was found.

2.3 Field-site model of flow route fluxes

The measured groundwater levels at the field-site were converted to field-site average groundwater depths, standard deviations of groundwater depths, and volumes of ponds on the soils surface to comply with the variables of the LGSI-model. The measured absolute groundwater levels at 31 locations within the field (Fig. 1c) were interpolated to a groundwater table for the entire field. Subsequently, this groundwater table was subtracted from a detailed DEM (5×5 m resolution) and all groundwater depths were grouped into a groundwater depth distribution. The volume of negative groundwater depths of this distribution quantifies the volume of ponds on the field. A mean and standard deviation of the groundwater depth distribution and the volume of ponds were calculated for all 57 weekly field-site groundwater depth surveys. Continuous groundwater level measurements in the 15 groundwater wells around the tube drain 1 (Fig. 1c) were used to interpolate between the weekly field average groundwater depths in order to create a continuous field average groundwater depth time series.

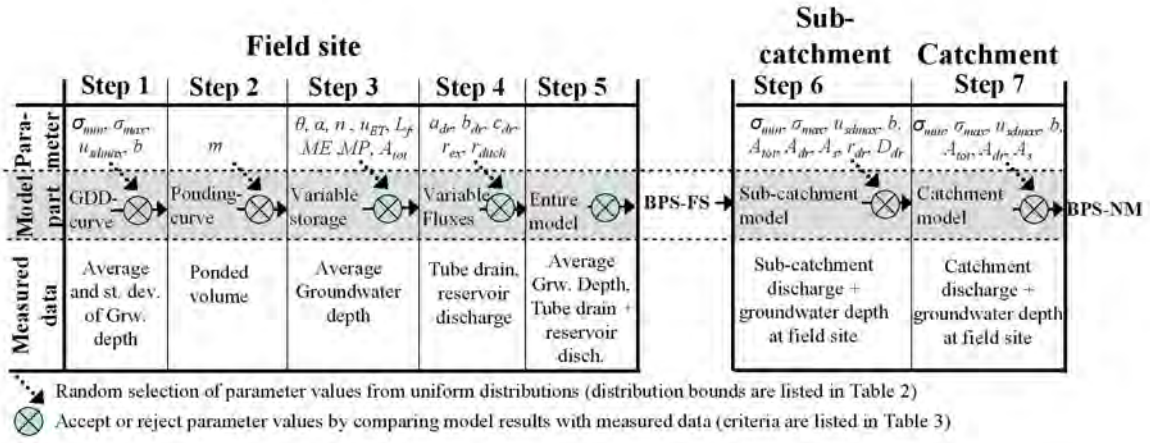

Fig. 2. Procedure for derivation of behavioral parameter sets for the fields-site and the nested-scales model. The first five steps yield behavioral parameter set for the field site (BPS-FS). All seven steps yield behavioral parameter sets for the nested scale model setup (BPS-N). The parameters are explained in Table 1.

The LGSI model of the field-site was conditioned on 5 sources of measured data:

- The measured relation between the average groundwater depth and the standards deviation of groundwater depth (GDD-data),
- The measured relation between the average groundwater depth and the volume of surface storage (Ponding-data),
- Time series of the spatial average groundwater depth,
- Time series of tube drain discharge,
- Time series of discharge measured by the in-stream reservoirs.

Figure 2 shows the five sequential steps that were followed to generate an ensemble of 500 Behavioral Parameter Sets (BPS-FS). The relations between model equations and measured data determined the specific order of these steps. The parameter distributions of BPS-FS were analyzed for parameter sensitivity and the model results were analyzed to quantify the uncertainty in flow route contributions to the total discharge owing to equifinality.

2.4 Catchment and nested scales model

The catchment model was conditioned on catchment discharge and a single groundwater level time series. The resulting ensemble of 500 Behavioral Parameter Sets is referred to by BPS-C. Table 2 lists the parameter ranges from which the parameter sets were generated and Table 3 list the criteria that devides the parameter space in behavioral and non-behavioral parameter sets. The ranges for the process-specific parameters were equal to those used for the field-site model. Only the ranges for the scale-specific parameters that described areas (i.e. the catchment area, the tube drained area, and the area of the surface water) were different.

The groundwater level time series, on which the catchment model was conditioned, was not considered representative for the dynamics of the groundwater storage of the entire catchment. From field experience, we estimated that at any one time, at least 10% of the catchment area had shallower groundwater and another 10% had deeper groundwater than the single observed level: the observed groundwater level thus was assumed to be within the 0.10 and 0.90 percentile but allowing for an average exceedence of 1.0 cm (GE, Table 3).

We attempted to constrain the uncertainty in flow route contributions of the catchment model by combining information from measurements from the field site, discharge measurements of a small sub-catchment, and discharge measurements at the catchment outlet in a nested-scales model setup. This nested-scales model consist of thee LGSI-models representing each of the scales: field-site, sub-catchment, and catchment. These models are connected by assuming that the parameters that describe the discharge response to groundwater depth are scale invariant (process-specific parameters, see Table 1), while the parameters that describe the spatial distribution of groundwater depths are assumed scale-specific (Table 1). The underlying hypothesis is that the differences between the observed hydrographs at the three scales are primarily an effect of a different spatial distribution of groundwater depths and resulting different active drainage areas.

All the seven steps of Fig. 2 were followed to derive an ensemble of 500 Behavioral Parameter Sets for the nested-scale

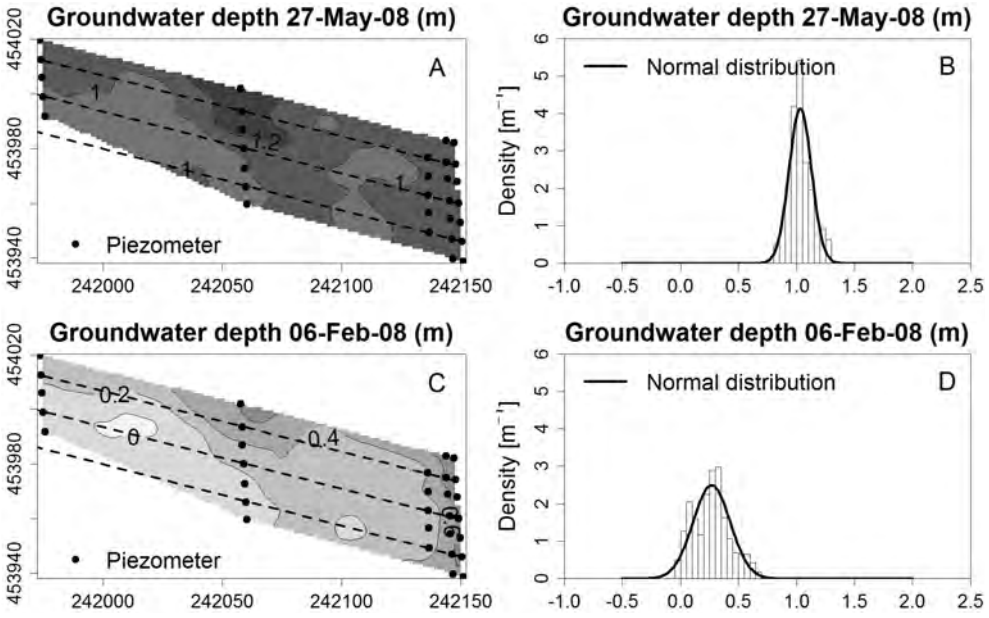

Fig. 3. Map of field-site groundwater depth for a dry **(A)** and a wet **(C)** day and the corresponding groundwater depth distributions (**B** and **D**).

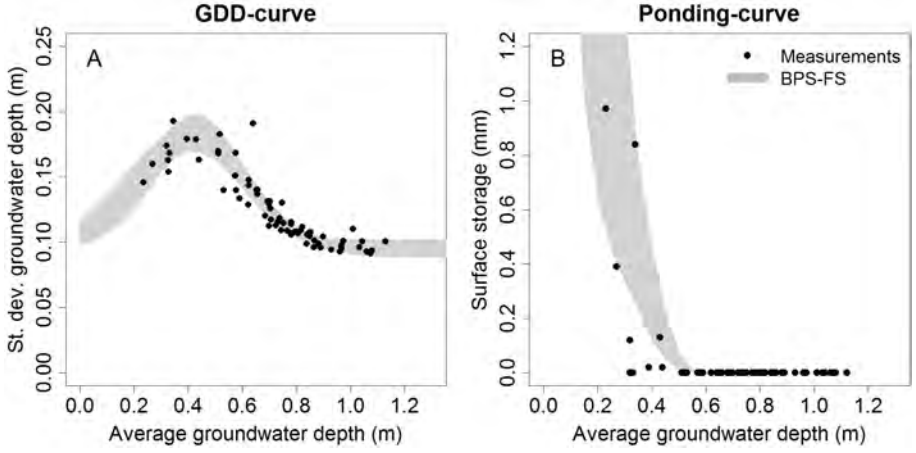

Fig. 4. Measured and simulated groundwater depth distribution curve (GDD-curve) **(A)** and Ponding curve **(B)** of the field-site. The grey area represents the ensemble of behavioral parameter sets of the field-site model (BPS-FS).

model setup (BPS-N). For the catchment-scales new constant drainage depth and drainage resistance were estimated, because the drainage depth and the time variant drainage resistance of the field site were specifically estimated for the three drains of the field site. Three drains probably do not represent the drain populations at larger scales. Hence, for the sub-catchment and catchment scale, we assumed a constant spatially averaged drainage resistance and drainage depth. The flow route contributions to discharge during the entire field-site monitoring period and the parameter distributions for BPS-C and BPS-N were compared to assess the added value of introducing nested-scale measurements.

2.5 Model validation

The BPS-C and the BPS-N were both validated for their ability to predict the catchment-scale discharge. For the validation we chose the period 1994–1995 for its high quality discharge data without data gaps and obvious measurement errors, and 1996–2001 for its episodes of extremely high discharges that are outside the discharge range of the calibration period.

The impact of uncertainty in flow route contributions to discharge on solute transport is demonstrated for the period of 16 March to 10 April 2008. For this period measured

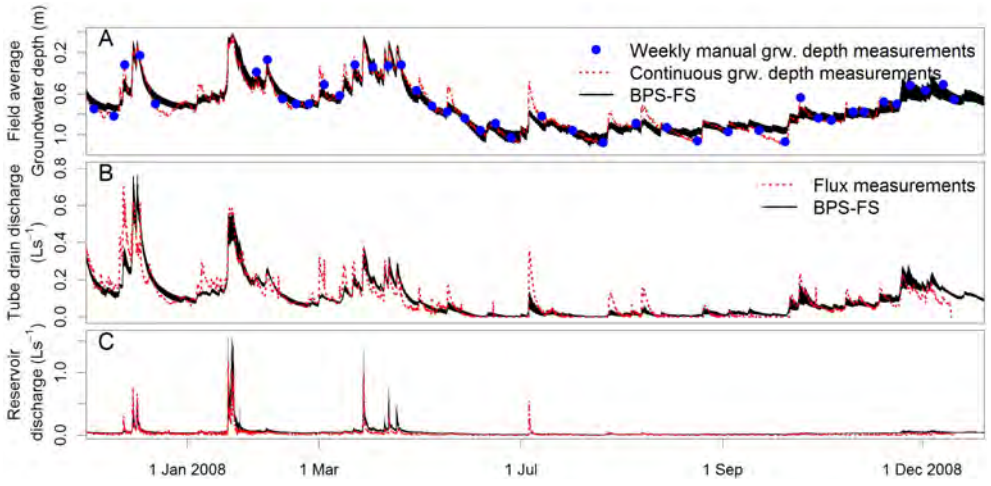

Fig. 5. Measurements and simulation results of the field site model (BPS-FS) for the spatially averaged groundwater depth **(A)**, tube drain discharge **(B)**, and reservoir discharge (combined flux of overland flow and groundwater flow, **C**) of the field-site. The black band gives the results of all behavioral parameter sets (BPS-FS).

nitrate loads are compared with loads calculated by both BPS-C and BPS-N. For each flow route constant concentrations were estimated from measurements reported in Van der Velde et al. (2010a): tube drain flow (72 mg L^{-1}), overland flow (9 mg L^{-1}) and direct rainfall (3 mg L^{-1}). The groundwater flow concentration was estimated to be 40 mg L^{-1}, based on the results of Rozemeijer et al. (2010c). Although constant nitrate concentration do not reflect the complexity of all processes affecting nitrate, Rozemeijer et al. (2010c) showed that tube drain nitrate concentrations remained relatively constant during this period and are not much affected by rainfall events. Also they showed that the spatial variability of nitrate concentrations is far larger than the temporal variation of individual flow routes at a single location. Therefore, constant flow route concentrations are a reasonable approximation for the one month period proposed for this comparison Note, that this approximation is only applied to illustrate the role of flow route fluxes in solute transport simulations and is not presented as a solute transport model.

3 Results and discussion

3.1 Field-site model results

Van der Velde et al. (2009) reported a decrease in the simulated groundwater depth variance as the groundwater depth increased in their catchment-scale groundwater model. For the field site this finding is corroborated by observations for a wet and a dry day (Fig. 3). Approximating the groundwater depth distributions by normal distributions introduced only small errors (Fig. 3). It shows that the extent of the features that dominate the shape of the water table such as tube drains, soil heterogeneity and micro topography are far smaller than the size of the field site. Consequently the central limit theorem guides the overall distribution of groundwater depths towards a normal distribution even for this single field site. Figure 4a shows the measured groundwater depth means and standard deviations for the 57 weekly groundwater depth surveys. The grey band in Fig. 4a represents the results of all behavioral parameter sets (BPS-FS) for the GDD-curve. This grey band is particularly narrow between average groundwater depths of 0.5 and 0.8 m (some of the data points are even outside it), indicating that the model results for storage and discharge are very sensitive to the GDD-curve in this range of average groundwater depths. GDD-curves outside the grey band, although they may closely fit the observed GDD-data, did not yield behavioral models for some of the other criteria such as storage or discharge. In Fig. 4b we plotted the measured ponding volumes and the ponding curves of all BPS-FS. Because the measured ponding volumes are relatively uncertain (they are difficult to measure and we have only a few measurements) we allowed for a larger curve error (CE, Table 3). This resulted in the grey band in Fig. 4b.

Figure 5a shows the measured and modeled spatially averaged groundwater depth. The model results are accurate, but some of the moderate groundwater level peaks are underestimated. This also caused an underestimation of the tube drain discharge (Fig. 5b) during these moderate groundwater level peaks. Overall, Fig. 5 shows that the LGSI-model is able to accurately describe the average groundwater depth, tube drain flow, and reservoir discharge (Eq. 2) simultaneously. All three time series were simulated with a Nash-Sutcliff (*NS*) coefficient (Nash and Sutcliff, 1970) exceeding 0.8.

A comparison of the LGSI-model results with the results of Rozemeijer et al. (2010b), who used the same dataset and a fully distributed HydroGeosphere (Therrien et al., 2009) model to simulate flow routes during a single discharge

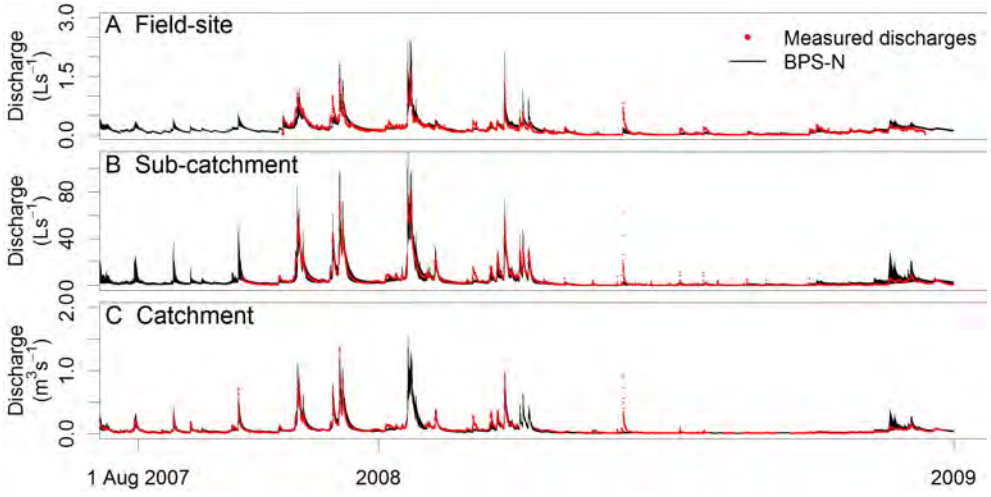

Fig. 6. Measurements and simulation results the nested-scales model for total discharges of the field-site **(A)**, sub-catchment **(B)** and entire catchment **(C)**. The black band gives the results of all behavioral parameter sets (BPS-N).

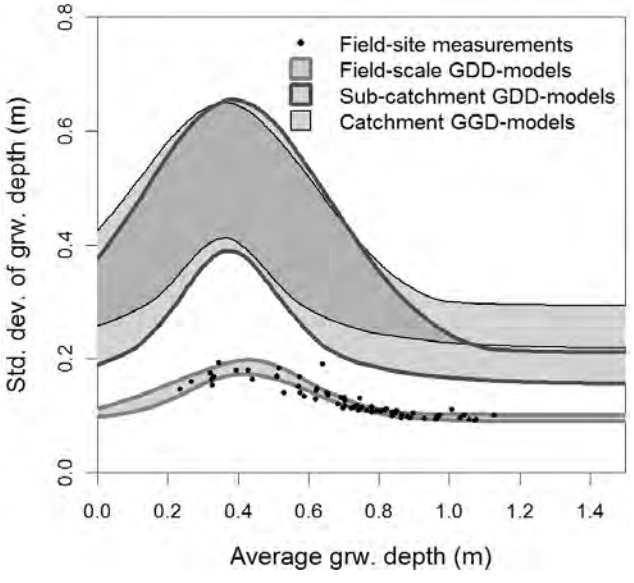

Fig. 7. Bandwidth of behavioral GDD-curves for the three scales of the nested-scales model (BPS-N). The dark grey area indicates overlap between ensembles of GDD-curves.

event, demonstrate that the relatively simple LGSI-model concepts can simulate the discharges of individual flow routes at least equally well as the HydroGeosphere model and hence constitute a very powerful tool for simulation and prediction of flow routes at a field site. Like Rozemeijer et al. (2010b), we found that measurements of both the storage of groundwater within the field and the corresponding discharge of flow routes are indispensible for an accurate model representation of the groundwater-surface water interaction.

3.2 Nested-scales model results

The discharges at all three nested-scales could be accurately described by parameter sets that share the same values for the process-specific parameters and differ only in scale-specific parameter values (Fig. 6). This result supports our hypothesis that scale effects in lowland hydrology can be attributed to scale differences in the shape of the groundwater table. These scale differences were quantified by the GDD-curves of the individual scales. Figure 7 shows the inferred ensemble of GDD-curves for the three nested scales. The differences between the field-site GDD-curve and the GDD-curve of the two catchment scales are much larger than those between the GDD-curves of both catchment scales. This is consistent with the small differences in the shape of the hydrographs between both catchment scales.

The individual flow route contributions (the median contribution of BPS-N) to discharge are shown in Fig. 8. The relative overland flow and groundwater flow contributions to discharge increase with increasing scale, at the expense of tube drain discharge. Table 4 gives the 10–90 percentile estimates of flow route contributions for the entire simulation period (this period equals the field-site measurement period). The uncertainty in the flow route contributions at the field site is constrained by many different types of measurements (see also Fig. 5). At the larger scales fewer measurements were available and consequently the uncertainty is much larger. In Table 4 we also compared the uncertainty of the flow route contributions calculated by BPS-N and BPS-C (the model constrained by catchment discharge only). The uncertainty in groundwater flow, overland flow, and direct rainfall is significantly reduced by introducing nested scale measurements. In contrast, the uncertainty of the tube drain discharge could hardly be reduced because the field-site tube drain depth and tube drain resistance could not be transferred to larger scales.

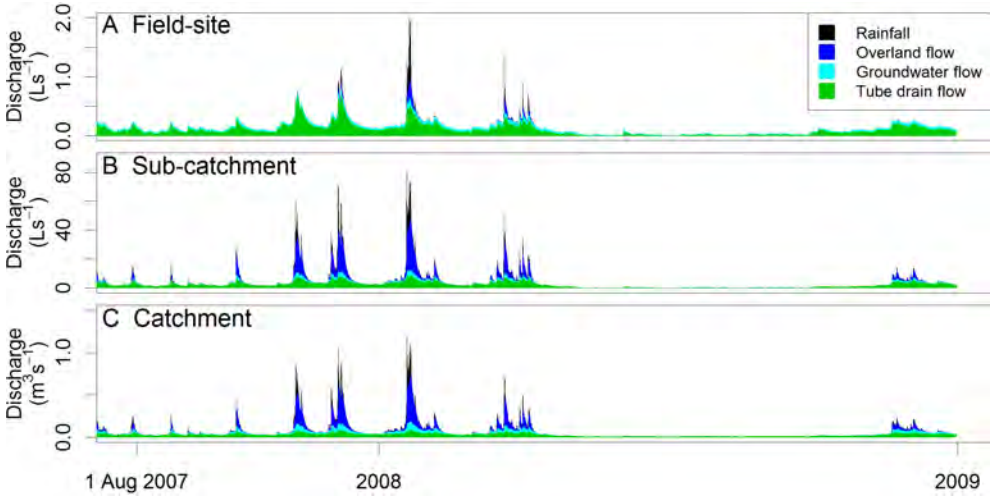

Fig. 8. Contributions of flow routes to total discharge. The displayed contribution is the median of contribution of all behavioral parameter sets (BPS-N).

Table 4. Calculated flow route contribution (0.1–0.9 quantiles) of BPS-N and BPS-C. The contribution is calculated over the period November 2007 through December 2008.

	BPS-N			BPS-C
	Field-site	Sub-catchment	Catchment	Catchment
Tube drain flow	0.76–0.79	0.34–0.61	0.25–0.50	0.21–0.51
Groundwater flow	0.10–0.15	0.06–0.16	0.12–0.27	0.14–0.50
Overland flow	0.04–0.07	0.24–0.42	0.27–0.41	0.18–0.37
Direct rainfall	0.03–0.05	0.07–0.11	0.08–0.11	0.03–0.10

We re-estimated the tube drain-specific parameters for the sub-catchment and catchment-scale by assuming them to be equal for both scales, but even this assumption did not reduce the uncertainty.

3.3 Parameter distributions

Figure 9 shows the distribution of the behavioral parameter sets of all process-specific parameters. Since the prior distributions were uniform, sensitive parameters are identified by markedly non-uniform distributions. For the field site model (BPS-FS) the most sensitive parameters are α, D_{dr}, a_{dr}, L_{f}, and r_{ditch}. The sensitivity of a_{dr} signals the importance of accounting for increasing tube drain resistance after tube drain cleaning. The lateral inflow of groundwater, L_{f}, which is the closing term for the water balance, could be determined accurately around 0.6 mm day^{-1}. Surprisingly insensitive parameters are θ_{s} and u_{ET}. The insensitivity of u_{ET} signals that evapotranspiration reduction at our relatively wet (high groundwater tables) field site might not be very important. As long as the value of u_{ET} is less than two standard deviations (20 cm) below the lowest average groundwater water table, the modeled evapotranspiration reduction is small and u_{ET} does not affect the model results.

At the catchment scale (BPS-C) on the other hand, u_{ET} is the most sensitive parameter. At this scale this parameters closes the overall water balance by increasing or reducing evapotranspiration (at the catchment scale there is not net lateral groundwater flow). Also m and r_{ex} are relatively sensitive as they directly control the discharge, which is the only calibration objective. All other parameters are insensitive and the uncertainty in flow route contribution is largely determined by the boundaries of the prior distributions.

For the nested-scales model (BPS-N) we sought process-specific parameter sets that can describe all three scales simultaneously. Figure 9 shows that the parameter distributions of BPS-N combine the constraints of the distribution of both BPS-FS and BPS-C. The added value of including nested-scale measurements to reduce parameter uncertainty is apparent for almost all process-specific parameters.

Figure 10 shows the distributions of the scale-specific parameters and again the reduction in parameter uncertainty by introducing nested-scale measurements is clear. Note that the distribution of σ_{min} tends to high values exceeding the

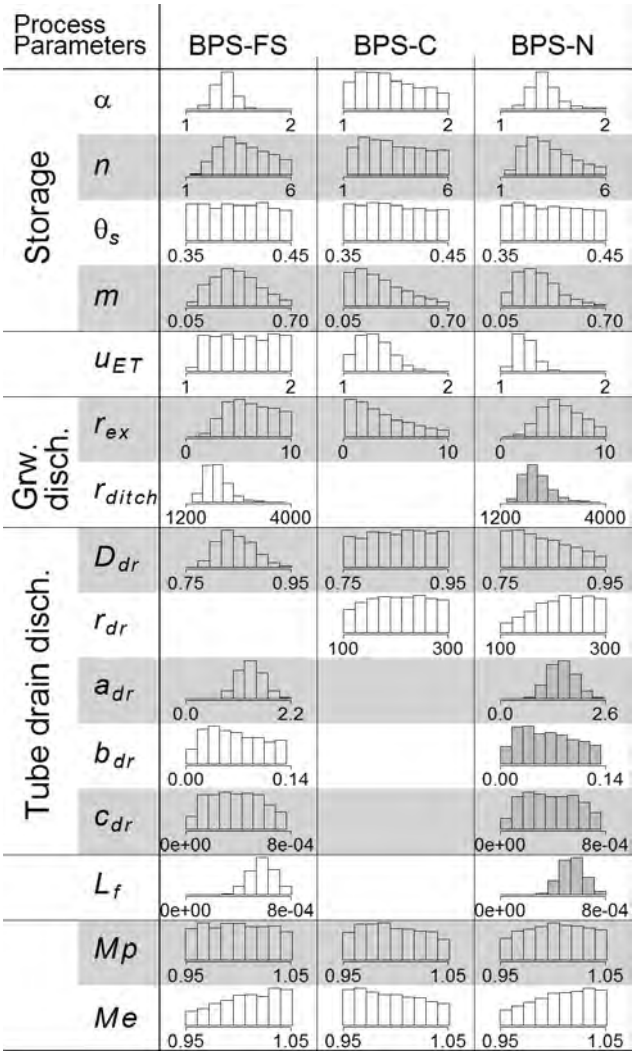

Fig. 9. Process-specific parameter distributions for the three behavioral parameter sets: BPS-FS, BPS-C, and BPS-N. The grey-filled distributions were only used for the field-site sub-model within the nested-scales model.

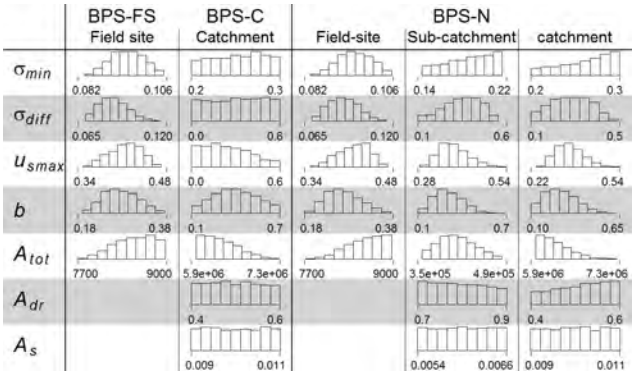

Fig. 10. Scale-specific parameter distributions for the three behavioral parameter sets: BPS-FS, BPS-C, and BPS-N.

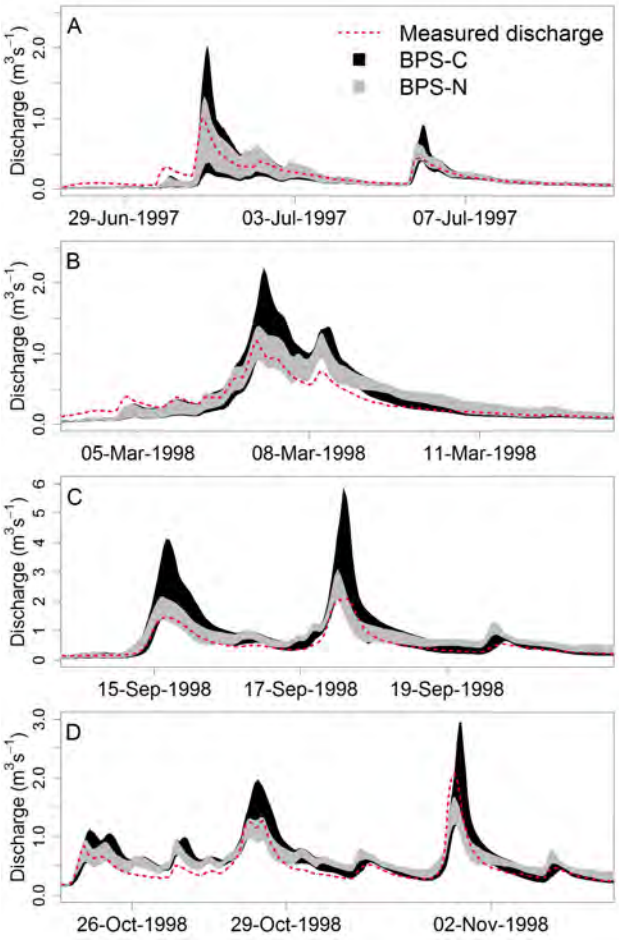

Fig. 11. Model validation results for extreme discharge events at the catchment outlet. The light grey band gives the results of the model calibrated on all nested-scale measurements, while the black band give the results of the model calibrated on catchment discharge only.

preset boundaries in Table 2 that were determined from a detailed DEM. Under dry conditions, the groundwater level is almost parallel to the soil surface. Variations in the groundwater depth under such conditions emanate largely from local variations in the soil surface elevation that are too small to appear in the phreatic level. However, under dry conditions a few very deep incisions of the stream produce most discharge. Around these incisions, the groundwater depth necessarily decreases sharply to zero at the stream bank, and these deviating groundwater depths produce outliers from the normal distribution valid for the rest of the catchment. As a consequence, the calibration tried to increase the groundwater depth variation (and thereby the range of $\sigma_{\min}$) under dry conditions to be able to generate low discharges. Because these discharges are low, the effect of underestimating discharges during dry conditions on the entire water balance of the catchment is small.

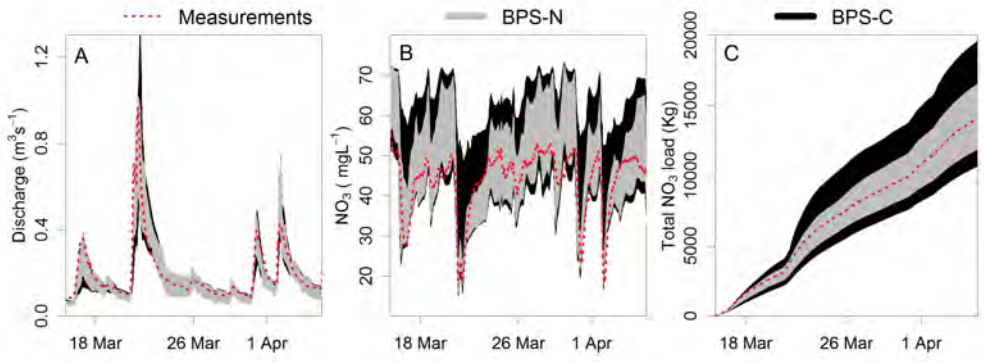

Fig. 12. Simulation results of BPS-N and BPS-C for discharge **(A)**, nitrate concentration via a flow route mixing analysis **(B)**, and cumulative nitrate load **(C)**.

In general, both Figs. 9 and 10 suggest that that the LGSI-model is highly over-parameterized when the model is only calibrated on discharge, because almost all parameters of BPS-C are insensitive (uniform distributions). However, nested-scale measurements and the assumption that scale differences originate only from differences in the shape of the groundwater table, increases the parameter sensitivity and allows most parameters to be conditioned.

3.4 Model validation

The BPS-N parameter sets performed slightly better than BPS-C for both validation periods. For the period 1994–1995 the BPS-N yielded an average *NS*-coefficient of 0.90, against 0.85 for the BPS-C. For the validation period 1996–2001 the average *NS*-coefficients were 0.79 for BPS-N and 0.73 for BPS-C. Although these may seem minor improvements in model performance when weighted against the efforts involved in the nested-scale monitoring, these model improvements are especially apparent in extreme discharge situations beyond the range of the calibration dataset. Figure 11 shows the validation results of the four most extreme discharge events during the period 1996–2001. From all events it is clear that the BPS-N much better predicted discharge than BPS-C (particularly the magnitude of the peaks), with far smaller uncertainty ranges. During the discharge event in Fig. 11c, which are the highest discharges measured in the past 20 years, the measured maximum discharge was more than twice the maximum discharge of the calibration period. During this period some of the BPS-C overestimated the discharge by a factor 3, while the BPS-N all predicted discharges close to the measured discharge.

The model improvement achieved by conditioning the model on nested-scale measurements is also apparent from the comparison between simulated and measured nitrate concentrations and nitrate loads (Fig. 12). There are hardly any differences between simulated discharges of BPS-N and BPS-C (Fig. 12a), as both parameter sets have been calibrated on this discharge time series. However, both parameter sets simulate different combinations of flow routes volumes that yield this overall discharge. The effects of flow route contributions on solute transport become apparent when each flow routes receives a unique concentration as demonstrated in Figs. 12b and c. Especially during peak discharges the spread of concentrations is much smaller for the parameter sets calibrated on the nested scale measurements (BPS-N). Hence, from the observed dynamics in nitrate concentrations it can be concluded that the flow route volumes of BPS-N are a more realistic representation of the flow process during high flow conditions. Figure 12c shows that conditioning on nested-scale data reduced the uncertainty in cumulative load estimates by roughly 50% and thus demonstrates the crucial role of accurate understanding of flow route discharges for predicting solute loads towards downstream surface water bodies.

4 Conclusions

Detailed and unique flux measurements at a pasture field site allowed us to formulate and calibrate our parsimonious LGSI-model. Even the very non-linear process of saturated overland flow was adequately simulated. An exceptional feature of this model is that the model concepts were designed around the available measurements (Van der Velde et al., 2009). Consequently, the parameters that describe the discharge and storage processes could all be conditioned on measurements. This yielded a field-site model that accurately described both storage and fluxes simultaneously.

The combined nested-scale measurement and model setup made it possible to combine discharge information of the field scale, a small sub-catchment, and the entire catchment. We demonstrated that the differences between hydrographs at the three scales could all be described by only changing the Groundwater Depth Distribution (GDD) curve, even though the hydrographs were markedly different. This

result supports our hypothesis that scale effects on discharge in lowland catchments are primarily an effect of differences in the spatial distributions of groundwater depths between different scales. Still, the range of GDD-curves that yielded good model results for the catchment scales was wide. This emphasizes the importance of spatially distributed groundwater depth monitoring to further condition these GDD-curves. This should lead to an even more solid foundation of the relation between storage and discharge, and hence to more reliable results.

Tube drain effluent is the most important route for nitrate towards the surface water network in lowland catchments. We were able to measure tube drain discharge at the field site and we concluded that almost 80% of the water (Van der Velde et al., 2010a) and 92% of the nitrate (Rozemeijer et al., 2010c) was transported by tube drains. It is by no means trivial to extrapolate these field-scale findings to the entire catchment. Our combined nested-scales observation and modeling approach could narrow down our estimates of the contribution of tube drain discharge to the discharge of a sub-catchment of 0.4 km^2 to 34–61% of the total discharge. For the entire catchment of 6.5 km^2, we estimated that 25–50% of the discharge originated from tube drains. These results not only demonstrate that we need to be careful extrapolating field experiment results to entire catchments but also show that nested-scale measurements are essential to understand and quantify the flow route contributions to the discharge of a catchment.

In this paper we demonstrated the potential of combined nested-scale monitoring and modeling for the Hupsel Brook catchment. However, many of our findings can be generalized. First of all, we showed that detailed field-site measurements of storage and flow routes provide the process-understanding that is needed to develop a model structure that adequately describes the catchment-specific flow routes. Secondly, we demonstrated that the combination of a relatively short period of nested-scale measurements with nested-scale models significantly constrains uncertainty in the contributions of groundwater flow, overland flow, and direct rainfall into surface waters. This reduction in flow route flux uncertainty significantly reduces uncertainty in nitrate load estimates as demonstrated by a simple flow route mixing approach. Finally, we showed that conditioning parameter sets on nested-scale measurements considerably improves discharge predictions compared to parameter sets constrained on discharge only. Model calibration on nested-scale measurements may not yield models with better calibration-statistics than models that are calibrated on catchment discharge alone, but the nested-scales model approach yields models that are able to predict (peak) discharges during validation periods more accurately. Improved quantifications and predictions of solute loads and peak discharges make the efforts involved in nested–scale monitoring worthwhile.

Appendix A

Notation

A1 Abbreviations

LGSI-model	Lowland Groundwater-Surface water Interaction model developed by Van der Velde et al. (2009).
GDD-curve	Groundwater Depth Distribution curve. Curve that relates the spatial standard deviation of the groundwater depth to its spatial average.
Ponding-curve	The relation between spatially averaged groundwater depth and the volume of ponds and surface waters.
BPS	Ensemble of 500 Behavioral Parameter Sets
BPS-FS, BPS-C, BPS-N :	BPS for each of the three models: field-site, catchment and nested-scales model, respectively.

A2 Symbols

A_{dr} [L^2]	Area within the catchment that is drained by subsurface tubes
A_s [L^2]	Area within the catchment covered with surface water
A_{tot} [L^2]	Catchment area
a_{dr} [–]*	Rate with which r_{dr} increases during wet periods
b_{dr} [T^{-1}]	Fractional rate with which r_{dr} decreases during dry periods
c_{dr} [L^3T^{-1}]	Threshold tube drain discharge: below this discharge r_{dr} decreases, above this discharge r_{dr} increases
CDE [–]	Cumulative discharge error, difference between cumulative measured and modeled discharge
CE [–]	Curve error, difference between measured data and the modeled GDD-curve or ponding-curve
D_{dr} [L]	Tube drain depth
e_{pot}[LT^{-1}]	Potential evapotranspiration
ET_{act} [LT^{-1}]	Actual evapotranspiration
E_Q[LT^{-1}]	Evaporation from ponded surface and surface waters
fu [L^{-1}]	Normal distribution function of groundwater depths
Fu [–]	Cumulative distribution function of groundwater depths
Fu^{-1}[–]	Inverse cumulative distribution function of groundwater depths

GE [L]	Average groundwater depth error between measured and modeled groundwater time series.
h[L]	Height above the groundwater table
L_f [LT^{-1}]	The constant lateral inflow of groundwater at the field site
m[–]	Fraction of groundwater levels above the soil surface that remain on the soil surface to constitute surface storage.
NS [–]	Nash-Sutcliff (Nash and Sutcliff, 1970) coefficient for time series.
P [LT^{-1}]	Rainfall
P_Q[LT^{-1}]	Rainfall on ponded surface or surface waters
Q [LT^{-1}]	Discharge at catchment outlet
Q_{grw} [LT^{-1}]	Groundwater flow for sub-catchment and entire catchment
$Q_{grw,field}$ [LT^{-1}]	Groundwater flow towards ditch at field site
Q_{ov} [LT^{-1}]	Overland flow
Q_{res}[LT^{-1}]	Discharge as measured by the in-stream reservoirs
r_{ditch} [T]	Resistance of the field site to groundwater flow towards the ditch.
r_{dr} [T]	Tube drain resistance
r_{ex}[T]	Groundwater exfiltration resistance
S_{sat} [L]	Storage in saturated zone normalized by area
S_{surf}[L]	Storage in ponds and surface water normalized by area
S_{unsat}[L]	Storage in unsaturated zone normalized by area
u[L]	Groundwater depth
$\langle u \rangle$[L]	Spatial average of the groundwater depth
u_{ditch} [L]	Depth of the field-site ditch relative to the mean surface elevation of the field site
u_{ET} [L]	Groundwater depth at which the actual evapotranpiration drops from e_{pot} to reduction depth
$u_{s\ max}$ [L]	Average groundwater depth at which the standard deviation of the groundwater depths is at its maximum
α[L], n[–]	Van Genuchten (1980) parameters that describe the soil water retention curve
σ_{diff} [L]	Maximum increase in the standard deviation of the groundwater depth
σ_{min} [L]	Minimal groundwater depth standard deviation
σ_u [L]	Groundwater depth standard deviation corresponding to a certain $\langle u \rangle$
θ_s [–]	Average porosity between highest and lowest groundwater table [–]

Appendix B

The overall water balance of the model normalized by area and zero lateral influx is given by:

$$\frac{\partial S_{sat}(t)}{\partial t} + \frac{\partial S_{unsat}(t)}{\partial t} + \frac{\partial S_{surf}(t)}{\partial t} = P(t) - ET_{act}(t) - Q(t) \quad \text{(B1)}$$

with saturated zone storage, S_{sat}[L], unsaturated zone storage, S_{unsat}[L], surface storage in streams, ditches and ponds, S_{surf}[L], the rainfall flux, P[LT^{-1}], actual evapotranspiration, ET_{act}[LT^{-1}], and discharge, Q[LT^{-1}]. The storage terms on the left-hand side of Eq. (A1) are described as a function of the distribution of groundwater depths, f_u[L^{-1}]. The change in saturated storage is expressed by the inverse of the change in unsaturated zone volume:

$$\frac{\partial S_{sat}(t)}{\partial t} = -\theta_s \frac{\partial}{\partial t}\left(\int_0^\infty f_u(t) u du\right) \quad \text{(B2)}$$

with the groundwater depth, u[L], the spatial averaged soil porosity, θ_s [–], the distribution in groundwater depths, f_u[L^{-1}], and the total unsaturated zone volume normalized by the catchment area $\int_0^\infty f_u(t) u du$. Note that the positive integration bounds implicate that the unsaturated zone does not exist for negative groundwater depths (ponding).

The volume of water stored in the unsaturated zone is described with a Van Genuchten (1980) relationship for soil moisture in an unsaturated zone at hydrostatic equilibrium. The change in unsaturated zone storage is described by:

$$\frac{\partial S_{unsat}(t)}{\partial t} = \theta_s \frac{\partial}{\partial t}\left(\int_0^\infty f_u(t) \int_0^u \left[1+(\alpha h)^n\right]^{\frac{1}{n}-1} dh du\right) \quad \text{(B3)}$$

The height above the groundwater level is denoted by h[L], and α [L^{-1}] and n [–] are the Van Genuchten parameters, with the residual volumetric water content equal to zero.

Water stored on the soil surface in ditches and ponds is described by a fixed fraction, m [–], of the total volume of groundwater heads above the soil surface:

$$S_{surf}(t) = -m \int_{-\infty}^0 f_u(t) u du \quad \text{(B4)}$$

The change in surface storage is given by:

$$\frac{\partial S_{surf}(t)}{\partial t} = -m \frac{\partial}{\partial t}\left(\int_{-\infty}^0 f_u(t) u du\right) \quad \text{(B5)}$$

Deep groundwater levels reduce the potential evapotranspiration. We chose a single cutoff level, u_{et}, below which no evapotranspiration is possible and above which potential

evapotranspiration, e_{pot}, occurs. This leads to the following expression for evapotranspiration:

$$\mathrm{ET}_{\mathrm{act}}(t) = e_{\mathrm{pot}}(t) \int_{-\infty}^{u_{\mathrm{et}}} f_u(t)du \tag{B6}$$

Furthermore, we assumed that a specific volume of groundwater is always stored in the same way (i.e., the moments of the distribution of groundwater depths only depend on the amount of storage and are not hysteretic). We are not interested in the exact configuration of storage within the catchment, and therefore assumed a normally distributed groundwater depth with an empirical relationship relating the standard deviation of groundwater depths,σ_{u} [L] to the spatial average of the groundwater depth $\langle u(t)\rangle$:

$$\sigma_{\mathrm{u}} = \sigma_{\mathrm{diff}} \cdot e^{-\left(\frac{\langle u(t)\rangle - u_{sd\max}}{b}\right)^2} + \sigma_{\min} \tag{B7}$$

Van der Velde et al. (2009) showed that this relation holds for field- and catchment-scales.

B1 Storage-discharge relationships

The contribution of specific flow routes to overall discharge largely determines the discharge quality. Therefore we subdivided the total discharge, Q, into four flow routes with distinctly different water chemistry:

$$Q(t) = Q_{\mathrm{drain}}(t) + Q_{\mathrm{grw}}(t) + Q_{\mathrm{ov}}(t) + P_{\mathrm{Q}}(t) - E_{\mathrm{Q}}(t) \tag{B8}$$

with Q_{drain} [LT^{-1}] groundwater discharge by tube drains, Q_{ov}[LT^{-1}], discharge by overland flow, and Q_{grw}[LT^{-1}], discharge of phreatic groundwater flow by ditch and stream drainage. Rain falling directly on the surface water network is denoted by P_{Q} [LT^{-1}] and evaporation from the surface water network is denoted by E_{Q}[LT^{-1}].

The tube drain discharge is calculated from the groundwater depth distribution by:

$$Q_{\mathrm{drain}}(t) = \frac{A_{dr}}{r_{dr}\, A_{\mathrm{tot}}} \int_{F_u^{-1}\left(A_{\mathrm{s}}/A_{\mathrm{tot}}\right)}^{D_{dr}} f_u(t) \cdot (D_{dr} - u)du \tag{B9}$$

with A_{dr} [L^2] the surface area occupied by tube drains, A_{tot} the catchment surface area, r_{dr} [T] the resistance of the soil to tube drain discharge and D_{dr} [T] the average depth of the tube drains. The fraction of catchment surface that is wet but has no tube drains, such as the surface area of ditches and streams, is denoted by A_{s}. This fraction is important under dry conditions when tube drainage stops and groundwater drainage by ditches and the stream takes over.

Van der Velde et al. (2009) made no distinction between overland flow and groundwater flow towards ditches and streams. They reasoned that the physical principles driving both fluxes are equal: groundwater level gradients driving groundwater from the soil into surface waters or ponds. In this study we follow the same line of reasoning but we want to separate both fluxes, because the two fluxes have distinctly different water quality. We hypothesize that under wet conditions, first all ditches start draining and only when the catchment becomes so wet that the drainage area, $F_u(0) = \int_{-\infty}^{0} f_u(t)du$, exceeds the surface area occupied by ditches, A_{s}, overland flow starts to occur. Now we can subdivide the groundwater flux into groundwater flow towards ditches and overland flow by the corresponding drainage area:

$$\begin{aligned} Q_{\mathrm{grw}} + Q_{\mathrm{ov}} &= \frac{m-1}{r_{\mathrm{ex}}} \int_{-\infty}^{0} f_u(t) \cdot u du \\ Q_{\mathrm{grw}} &= \frac{m-1}{r_{\mathrm{ex}}} \int_{-\infty}^{\min\left(F_u^{-1}\left(\frac{A_{\mathrm{s}}}{A}\right),0\right)} f_u(t) \cdot u du \\ Q_{\mathrm{ov}} &= \frac{m-1}{r_{\mathrm{ex}}} \int_{\min\left(F_u^{-1}\left(\frac{A_{\mathrm{s}}}{A}\right),0\right)}^{0} f_u(t) \cdot u du \end{aligned} \tag{B10}$$

with r_{ex} [T] the resistance of the soil to groundwater flow towards surface water and ponds. The term $\min\left(F_u^{-1}\left(\frac{A_{\mathrm{s}}}{A}\right),0\right)$ divides the negative part of the distribution of groundwater depth f_u(i.e. areas with ponding) in two areas: an area with groundwater flow and an area with overland flow, where $F_u^{-1}\left(\frac{A_{\mathrm{s}}}{A}\right) < 0$.

The amount of rain that falls on the active drainage area and is discharged immediately is:

$$P_{\mathrm{Q}} = \langle p\rangle \int_{-\infty}^{0} f_u(t)du \tag{B11}$$

The evaporation of surface water is:

$$E_{\mathrm{Q}} = \left\langle e_{\mathrm{pot}}(t)\right\rangle \int_{-\infty}^{0} f_u(t)du \tag{B12}$$

Appendix C

C1 Model error expressions

Four error-terms divide the parameters space in behavioral and non-behavioral parameter sets. The Curve Error term quantifies the average normalized distance between the measurements and the modeled GDD-curve or Ponding curve:

$$\mathrm{CE} = \frac{1}{n}\sum_{i=1}^{n} \min\left(\sqrt{\left(\frac{\langle u\rangle_i - \langle u\rangle_c}{\overline{\langle u\rangle}}\right)^2 + \left(\frac{y_i - y_c}{\overline{y}}\right)^2}\right),$$
$$\text{with } y_c = f(\langle u\rangle_c) \tag{C1}$$

The minimum function identifies the minimum normalized distance between the measurements and the model-curve. The number of measurements is denoted by n; $\langle u \rangle_i$ is the spatially averaged groundwater depth at the time measurement i was obtained; $\langle u \rangle_c$ is an spatially averaged groundwater depth defined by the GDD-curve or the Ponding-curve; $\overline{\langle u \rangle}$ is the temporal average of the measured spatially averaged groundwater depth (i.e., the groundwater depth averaged over space and time); y_i is the measured variable (the standard deviation of the groundwater depth for the GDD-curve, and the ponding volume for the Ponding-curve); y_c is the same variable defined by the model curve, and $\overline{y}$ the temporal average of the measured values of this variable.

The second error term quantifies the error in the total water balance:

$$\mathrm{CDE} = \left| 1 - \frac{\sum_{i=1}^{n} Q_{\mathrm{mod},i}}{\sum_{i=1}^{n} Q_{\mathrm{meas},i}} \right| \tag{C2}$$

With $Q_{\mathrm{mod},i}$ [$\mathrm{L^3T^{-1}}$], the modeled discharge corresponding to measured discharge Q_{meas}, i [$\mathrm{L^3T^{-1}}$]. The dynamics of groundwater depth and discharge time series (both defined by the variable V in the following equation) are evaluated by the Nash-Sutcliff statistic (Nash and Sutcliffe, 1970)

$$NS = 1 - \frac{\sum_{i=1}^{n} \left(V_{\mathrm{meas},i} - V_{\mathrm{mod},i} \right)^2}{\sum_{i=1}^{n} \left(V_{\mathrm{meas},i} - \overline{V_{\mathrm{meas}}} \right)^2} \tag{C3}$$

with $\overline{V_{\mathrm{meas}}}$ [$\mathrm{L^3T^{-1}}$] the average measured discharge or groundwater depth.

The simulated groundwater dynamics for both catchment scales are compared with a single measured groundwater level time series. This comparison is evaluated by:

$$\mathrm{GE} = \frac{1}{n} \sum_{i=1}^{n} f\mathrm{GE}_i \tag{C4}$$

with fGE [–] a measure of the degree to which a measured groundwater depth, $u_{\mathrm{meas},i}$[L], is outside the acceptable bounds ($Ul_{\mathrm{min},i}$ [L] and $Ul_{\mathrm{max},i}$ [L]) of the modeled groundwater depth distribution:

$$f\mathrm{GE}_i = u_{\mathrm{meas},i} - Ul_{\mathrm{max},i} \text{ if } u_{\mathrm{meas},i} > Ul_{\mathrm{max},i}$$

$$f\mathrm{GE}_i = Ul_{\mathrm{min},i} - u_{\mathrm{meas},i} \text{ if } u_{\mathrm{meas},i} < Ul_{\mathrm{min},i}$$

$$f\mathrm{GE}_i = 0 \text{ if } Ul_{\mathrm{min},i} < u_{\mathrm{meas},i} < Ul_{\mathrm{max},i}$$

The acceptable bounds, Ul_{min} and Ul_{max}, are a fixed quantile of the modeled distribution and are recalculated for each time step, i, based on an estimate of the representativity of a location where groundwater depths are measured for a certain area. We estimated that the measured average groundwater depth at the field site should always be within the 0.20–0.80 percentile of all groundwater depths in the subcatchment and within 0.1–0.90 of the groundwater depths of the entire catchment.

Acknowledgements. The authors wish to acknowledge Deltares and Alterra for funding this study. This study would not have been possible without the data collection by Piet Warmerdam, Jacques Kole, KNMI, and waterboard Rijn and Ijssel. The help of technicians Pieter Hazenberg and Harm Gooren to develop the field-site measurement setup is highly appreciated.

Edited by: A. Butturini

References

Appels, W. M., Bogaart, P. W., and Van der Zee, S. E. A. T. M.: Influence of spatial variations of microtopography and infiltration on surface runoff and field scale hydrological connectivity, Ad. Water Resour., 34, 303–313, 2011.

Bentley, W. J. and Skaggs, R. W.: Changes in entrance resistance of subsurface drains, J. Irrig. Drain E-Asce, 119, 584–598, 1993.

Beven, K. J. and Freer, J. E.: Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology, J. Hydrol., 249, 11–29, 2001.

Botter, G., Milan, E., Bertuzzo, E., Zanardo, S., and Rinaldo, A.: Inference from catchment-scale tracer circulation experiments, J. Hydrol., 369, 368–380, 2009.

Clark, M. P., Rupp, D. E., Woods, R. A., Tromp-van Meerveld, H. J., Peters, N. E., and Freer, J. E.: Consistency between hydrological models and field observations: linking processes at the hillslope scale to hydrological responses at the watershed scale, Hydrol. Process., 23, 311–319, 2009.

De Vos, J. A., Hesterberg, D., and Raats, P. A. C.: Nitrate leaching in a tile-drained silt loam soil, Soil Sci. Soc. Am. J., 64, 517–527, 2000.

Didszun, J. and Uhlenbrook, S.: Scaling of dominant runoff generation processes: Nested catchments approach using multiple tracers, Water Resour. Res., 44, doi:10.1029/2006WR005242, 2008.

Kirchner, J. W.: Getting the right answers for the right reasons: Linking measurements, analyses, and models to advance the science of hydrology. Water Resour. Res., 42, W03S04, doi:10.1029/2005WR004362, 2006.

Jencso, K. G., McGlynn, B. L., Gooseff, M. N., Wondzell, S. M., Bencala, K. E., and Marshall, L. A.: Hydrologic connectivity between landscapes and streams: Transferring reach- and plot-scale understanding to the catchment scale, Water Resour. Res., 45, doi:10.1029/2008WR007225, 2009.

Makkink, G. F.: Testing the Penman formula by means of lysimeters, J. Int. Water Eng., 11, 277–288, 1957.

Nangia, V., Gowda, P. H., Mulla, D. J., and San, G. R.: Modeling impacts of tile drain spacing and depth on nitrate-nitrogen losses, Vadose zone J., 3, 61–72, 2010.

Nash, J. E. and Sutcliffe, J. V.: River Flow Forecasting through Conceptual Models, Part I-A, Discussion of Principles, J. Hydrol., 10, 282–290, 1970.

Oenema, O., Oudendag, D., and Velthof, G. L.: Nutrient losses from manure management in the European Union, Livest. Sci., 112, 261–272, 2007.

Rodgers, P., Soulsby, C., Waldron, S., and Tetzlaff, D.: Using stable isotope tracers to assess hydrological flow paths, residence times and landscape influences in a nested mesoscale catchment, Hydrol. Earth Syst. Sci., 9, 139–155, doi:10.5194/hess-9-139-2005, 2005.

Rozemeijer, J. C., Van der Velde, Y., De Jonge, H., Van Geer, F. C., Broers, H. P., and Bierkens, M. P. F.: Application and evaluation of a new passive sampler for measuring average solute concentrations in a catchment-scale water quality monitoring study, Environ. Sci. Technol., 44, 1353–1359, 2010a.

Rozemeijer, J. C., Van der Velde, Y., McLaren, R. G., Van Geer, F. C., Broers, H. P., and Bierkens, M. F. P.: Using field scale measurements of flow route contributions to improve integrated model representations of dynamic groundwater-surface water interactions, Water Resour. Res., 46, W11537, doi:10.1029/2010WR009155, 2010b.

Rozemeijer, J. C., Van der Velde, Y., Van Geer, F. C., Broers, H. P., and Bierkens, M. F. P.: Direct quantification of the tile drain and groundwater flow route contributions to surface water contamination: from field-scale concentration patterns in groundwater to catchment-scale surface water quality, Environ. Pollut., 158, 3571–3579, 2010c.

Shaman, J., Stieglitz, M., and Burns, D.: Are big basins just the sum of small catchments?, Hydrol. Process., 18, 3195–3206, 2004.

Sivapalan, M.: Process complexity at hillslope scale, process simplicity at the watershed scale: is there a connection?, Hydrol. Process. 17, 1037–1041, 2003.

Soulsby, C., Tetzlaff, D., Dunn, S.M., and Waldron, S.: Scaling up and out in runoff process understanding: insights from nested experimental catchment studies, Hydrol. Process., 20, 2461–2465, 2006.

Tetzlaff, D., Waldron, S., Brewer, M. J., and Soulsby, C.: Assessing nested hydrological and hydrochemical behaviour of a mesoscale catchment using continuous tracer data, J. Hydrol., 336, 430–443, 2007.

Tetzlaff, D., McDonnell, J. J., Uhlenbrook, S., McGuire, K. J., Bogaart, P., Naef, F., Baird, A., Dunn, S. M., and Soulsby, C.: Conceptualizing catchment processes: simply complex?, Hydrol. Process., 22, 1727–1730, 2008.

Therrien, R., McLaren, R. G., Sudicky, E. A., and Panday, S. M.: HydroGeoSphere, A Three-dimensional Numerical Model Describing Fully-integrated Subsurface and Surface Flow and Solute Transport (Draft), Groundwater Simulations Group, University of Waterloo, Waterloo, Canada, 2009.

Tiemeyer, B., Kahle, P., and Lennartz, B.: Nutrient losses from artificially drained catchments in North-Eastern Germany at different scales, Agr. Water Manage., 85, 47–57, 2006.

Tiemeyer, B., Lennartz, B., and Kahle, P.: Analyzing nitrate losses from an artificially drained lowland catchment (North-Eastern Germany) with a mixing model, Agr. Ecosyst. Environ., 123, 125–136, 2008.

Tiemeyer, B., Kahle, P., and Lennartz, B.: Designing monitoring programs for artificially drained catchments, Vadose Zone J., 9, 14–24, 2010.

Uchida, T., Asano, Y., Onda, Y., and Miyata, S.: Are headwaters just the sum of hillslopes?, Hydrol. Process. 19, 3251–3261, 2005.

Van der Molen, D. T., Portielje, R., De Nobel, W. T., and Boers, P. C. M.: Nitrogen in Dutch freshwater lakes: trends and targets, Environ. Pollut., 102, 553–557, 1998.

van der Velde, Y., de Rooij, G. H., and Torfs, P. J. J. F.: Catchment-scale non-linear groundwater-surface water interactions in densely drained lowland catchments, Hydrol. Earth Syst. Sci., 13, 1867–1885, doi:10.5194/hess-13-1867-2009, 2009.

Van der Velde, Y., Rozemeijer, J. C., De Rooij, G. H., Van Geer, F. C., and Broers, H. P.: Field-scale measurements for separation of catchment discharge into flow route contributions, Vadose Zone J., 9, 25–35, 2010a.

Van der Velde, Y., De Rooij, G. H., Rozemeijer, J. C., Van Geer, F. C., and Broers, H. P.: The nitrate response of a lowland catchment: on the relation between stream concentration and travel time distribution dynamics, Water Resour. Res., 46, W11534, doi:10.1029/2010WR009105, 2010b.

Van Genuchten, M. Th.: A closed-form equation for predicting the hydraulic conductivity of unsaturated soils, Soil Sci. Soc. Am. J., 44, 892–898, 1980.

Van Ommen, H. C., Dijksma, R., Hendrickx, J. M. H., Dekker, L. W., Hulshof, J. , and Van Den Heuvel, M.: Experimental assessment of preferential flow paths in a field soil, J. Hydrol., 105 253–262, 1989.

Vitousek, P. M., Naylor, R., Crews, T., David, M. B., Drinkwater, L. E., Holland, E., Johnes, P. J., Katzenberger, J., Martinelli, L. A., Matson, P. A., Nziguheba, G., Ojima, D., Palm, C. A., Robertson, G. P., Sanchez, P. A., Townsend, A. R., and Zhang, F. S.: Nutrient Imbalances in Agricultural Development, Science, 324, 1519–1520, 2009.

Wösten, J. H. M., Bouma, J., and Stoffelsen, G. H.: Use of soil survey data for regional soil water simulation models, Soil Sci. Soc. Am. J., 49, 1238–1244, 1985.

Wösten, J. H. M., Veerman, G. J., De Groot, W. J. M., and Stolte, J.: Waterretentie- en doorlatendheidskarakteristieken van boven- en ondergronden in Nederland: de Staringreeks. Vernieuwde uitgave; 86 pp., http://www2.alterra.wur.nl/Webdocs/PDFFiles/Alterrarapporten/AlterraRapport153.pdf, 2001.